Why Johnny Can’t Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts

Year: 2023 | Citations: 881 | Authors: J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, Qian Yang

---

Abstract

Pre-trained large language models (“LLMs”) like GPT-3 can engage in fluent, multi-turn instruction-taking out-of-the-box, making them attractive materials for designing natural language interactions. Using natural language to steer LLM outputs (“prompting”) has emerged as an important design technique potentially accessible to non-AI-experts. Crafting effective prompts can be challenging, however, and prompt-based interactions are brittle. Here, we explore whether non-AI-experts can successfully engage in “end-user prompt engineering” using a design probe—a prototype LLM-based chatbot design tool supporting development and systematic evaluation of prompting strategies. Ultimately, our probe participants explored prompt designs opportunistically, not systematically, and struggled in ways echoing end-user programming systems and interactive machine learning systems. Expectations stemming from human-to-human instructional experiences, and a tendency to overgeneralize, were barriers to effective prompt design. These findings have implications for non-AI-expert-facing LLM-based tool design and for improving LLM-and-prompt literacy among programmers and the public, and present opportunities for further research.